BOOK REVIEWS.

Elements of Statistics. By ARTHUR L. BOWLEY. London: P. S. King & Son, 1901. 8vo, pp. vii+330.

MR. BOWLEY is at present lecturer in statistics at the London School of Economics and Political Science, but this does not go far to describe his brilliant work in statistical studies. He has taken both the Cobden and Adam Smith prizes at Cambridge, as well as the Guy silver medal of the Royal Statistical Society, and was Newmarch lecturer in 1897–8 at University College, London. His study of wages in the United Kingdom in the nineteenth century contained examples of the application of his methods, and in the book before us he has undertaken to give a more full and better rounded account of statistical methods as they have been developed under the teachings of the mathematical school. His experience and happy results call for consideration and careful study, although his exposition as a whole may not be accepted as the best manner of illustrating the practical application of statistical methods.

It may be said at the start the book is a forbidding one, not to be hastily read, or lightly laid down, but calling for an amount of study and some knowledge of higher mathematics denied to the majority of those who have done excellent work in statistics. The entrance of the more refined methods of the pure mathematician into what passed for the province of statistics was inevitable, as so many problems in statistics admit of the application of the doctrine of chance or probability. What Quetelet began tentatively has been developed, and it is only necessary to name Edgeworth, Pearson, Galton, and Yule to show how far the study has progressed in England. The formal instruction of this school is of recent origin, as the almost pathetic note of Mr. Yule records:

> In the following session, however, (1894–5) Professor Pearson started lecturing on the theory of statistics at University College, giving two hours a week to a small but enthusiastic class of two students, Miss Alice Lee, demonstrator in physics at the Bedford College, and myself. The course covered the theory of normal and skew variation, dissection of compound normal curves, and theory of normal correlation.

The very enumeration is enough to show the new factors injected into the study of statistics by the advent of the mathematical statist.

This new spirit of interpretation has been mastered by Mr. Bowley, and his book may thus be accepted as the latest and best summary of its methods. He shows a familiarity with the results of the studies of Pearson and Edgeworth; he has known the use made of some of them by Wilson Fox in his report on the "Wages and Earnings of Agricultural Labourers," and he has applied his own knowledge in his study of English wages in the last century. Both in theory and in practice Mr. Bowley was well equipped to undertake his task, and it only remains to examine how far he has been successful in meeting the requirements of those who would understand the principles of statistical methods. In only one instance has he devoted a chapter to a special subject — the index number — and solely for the reason that a good opportunity was offered for illustrating the principles laid down in the section on "accuracy." Apart from this instance he has restricted his study to methods, thus gaining in consistency and logical arrangement, but at the expense at times of clearness or ready comprehension. To apply each principle to a number of examples, or to illustrate each formula by a number of differing problems would have expanded the volume to an undue bulk. Yet it could be wished that more examples had been solved, and a larger number of applications made, as the beginner must be led by slow stages, with each stage definitely marked for him.

The chapters devoted to the scope and meaning of statistics, and general method of statistical investigation need not be examined, as they contain what is to be found in the best treatises on statistics. The application of methods to the population census, the wage census, the work of the labor department, and to the statistics of England's foreign trade, is well developed and the examples are happily chosen. Of somewhat mechanical importance is the section on tabulation, where the experience of the United States has been wider, and that of Europe more varied still, leading to an impression that the continental statisticians are neglecting the broad foundations of the "science," to amuse themselves and astonish others by a reckless generalization from imperfect data. The results, however interesting and suggestive, remain beyond the field of true science, and their worth is of too doubtful a character to meet with a ready acceptance.

It is at this point that Mr. Bowley enters upon the investigations

which have been peculiar to the mathematical statistician, and his chapters on averages, accuracy, interpolation, and the application of the theory of probability to statistics, constitute his most positive contributions to the teaching of statistics.

The most useful instrument in statistical investigation is the average or mean, for the words are in ordinary usage synonymous. Mr. Bowley suggests a distinction by regarding an average as a purely arithmetical conception, while a mean must imply some objective quantity. The average length of life in a varied population would answer to the first description, while Quetelet's "average man" would answer to the second. Such a distinction avoids the necessity urged by some writers of looking upon any average as fictitious, a mere convenience in measuring or estimating a number of differing elements. The uses to which averages may be put are many, but it is difficult to decide at times what form the average should take, and it is because this question is not often fully considered that so much loose statistical writing is to be encountered. We all know the free resort to averages, percentages, per capita, and similar forms of means or averages in the current discussions of economic problems, and how seldom the results are based upon true values or upon correct methods. An average is imposing, and is simple, unencumbered by the perplexities of the details from which it has been figured; but no average ought to be accepted without some knowledge of these details, for nothing is simpler than to give a result due to bias, prejudice, or miscalculation.

The simple arithmetical average, obtained by taking the sum of a number of elements and dividing by the number of elements, is that which is most generally employed, and fully meets most of the ordinary conditions to be encountered. Yet its range is more limited than is usually believed, and more delicate measurements must be obtained, measurements in which due allowance is given to the "weight" of each of the elements included in the computation. In estimating, for example, the average wages received in a certain district, the number of men receiving each rate of wage is important; or, if the estimate is made by towns, an unusually high wage paid in a very unimportant locality may have the effect of giving too high an average for the whole. The true method is to take the wages of all the laborers in the district and divide the sum by the number of laborers. Even this operation will give an approximate result, but the conditions involved, that all the wages and the full number of laborers are known, can

never be fulfilled in practice. It is usual, therefore, and the results justify the method, to "weight" the elements according to their importance. In forming an index-number of the change in prices this weighting is essential, and Mr. Bowley has reproduced the basis of the index-number recommended by the committee appointed by the economic section of the British association in 1888 (p. 225.) There is, however, no agreement upon the proper weights to be used. In spite of the many discussions upon this matter, Mr. Bowley does not attach much importance to it, for "one of the most convenient facts of statistical theory is that, given certain conditions, the same result is obtained whatever logical system of weights is applied." Two examples of different weightings are given by our author, both of them based upon the wage census (pp. 113–118), and should be carefully studied. We can only quote his conclusion, because it contains a needed warning:

Since the true system of weights which would reduce the general average to our definition must be allied to some of those here adopted, and can hardly show greater divergence from 12*s.* 4*d.* than these do, we may feel confident that the true average is within, say, 3*d.* of this figure. The original items varied from 8*s.* 6*d.* to 19*s.*; the averages, even those based on the most extravagant methods, are contained by the limits 12*s.* and 13*s.* 1¾*d.* Without some such agreement as this, we should have no clue to the magnitude of the error introduced by erroneous weights. This is of the greatest importance, because in many statistical questions the true weights are undefinable or incalculable; now it is seen that, given certain conditions, there is no need to calculate or define the weights. Notice, however, that no system of weights can remove an original bias common to all the figures; if, for example, winter wages throughout were 1*s.* less than here reckoned, the corresponding deficit would appear unchanged in all the averages found. So we arrive at a very important precept: *in calculating averages give all your care to making the items free from bias, and leave the weights to take care of themselves* (pp. 117–118).

Passing from the arithmetic average, Mr. Bowley takes up the "mode," which is the "position of greatest density, or the position of the maximum ordinate," or a predominant rate. One is tempted to enter a protest against the introduction of purely scientific terms into a subject regarded as popular, and to rebel against the manifest intention to force into statistics such terms as median, quartile, decile, and percentile, terms borrowed, we believe, from astrology, and therefore infected with the flavor of a pseudo-scientific jargon. A little reflection

will show, however, that if these terms are useful in mathematics, they must be used in mathematical statistics, and the mode has already proved its utility. The arithmetic average has its limitations; the mode takes us one step farther. In dealing with a large number of different elements, with wide divergences, the mode determines the position of maximum importance. The illustrations given by Mr. Bowley are happily chosen, and his discussion of the proper location of the mode when once found in a general manner is admirably clear, and becomes convincing when developed graphically (see pp. 121 and 154). While the arithmetic average is affected by extremes, and thus misleading unless properly weighted, the mode is not influenced by them in the least, for they are outside of the region of maximum density. There is one limitation to the mode, that it must indicate a type, on either side of which the divergences fall. Unless this type is present the method can hardly be applied with success. "No single measurement expresses completely even the economic condition of a group of workmen, but if we are taking a single measurement, that of the 'mode' is often the most useful. It is at the mode that we find the greatest number of whose greatest good we may be thinking. Whereas the arithmetic mean and the 'median' may correspond to no reality, but be merely numerical conceptions, the mode is precisely that number for which most instances can be found. It shows the commonest result, that most often obtained, and is of very general application."

The position of the mode is often indefinite, and the median — the half-way number or interval in a list of items arranged in ascending order of magnitude — is more advantageous, and may be found even when the available information on the items in the list is neither accurate nor complete. Its simplicity is a strong argument in its favor, and it is not affected by exceptional entries. "There is yet another advantage, perhaps more important, that the median is applicable to quantities which are not capable of measurement at all." This development, due to Francis Galton, might be supposed to make the median of universal applicability, but a disadvantage intervenes to limit its usefulness. "The median of a series of observations may be totally removed from its type, and, in fact, may not be situated near any of the different objects which are observed. . . . .The median is then chiefly useful when we are dealing with a series of objects of which the main parts lie fairly close together; a few extremes do not affect it."

After dismissing the geometric average in a short paragraph as not a very useful instrument, Mr. Bowley closes his discussion of the average with an attempt to discriminate the employment of the different forms:

> Different methods will apply to groups of various classes; each must be taken on its own merits. A good and suitable average has the following characteristics: If there is a type, it shows it; it gives due influence to extreme cases; it is not easily affected by errors or much displaced by slight alterations in systems of calculation; and it is easily calculated. The relative positions of the different kinds of averages dealt with gives some information as to the general nature of the group to which they refer. The arithmetic, median, and mode are close together, if the group is symmetrical. The arithmetic average is probably above the median, if we have a small group at a high degree. The arithmetic average is generally below the median, if there is an absence of high numbers, and concentration a little above the average. The mode will be badly defined, if our group is not homogeneous. The mode will probably be below the arithmetic average, if there is a small group at a high degree. The mode is well marked, if the distribution is uniform. These rules are only tentative and easily nullified by exceptional circumstances.

We have thus dwelt upon Mr. Bowley's treatment of averages because it is characteristic of his general method. Beginning with the simple form he has passed to the more complex, translating them into mathematical terms, and clothing them with the universality of mathematics—subject however to the imperfections of existent knowledge and the many gaps to be found in statistical records. So far as the more popular methods are involved the new language of algebraic or geometrical expression need not confuse, and it does serve to point out the readiness with which old formulæ may be applied to novel problems, producing interesting results. But is it not too early to look upon statistics as a science capable of mathematical exactitude and absoluteness? In dealing with "interpolation" the many formulæ possible are recorded, yet Mr. Bowley admits that "It cannot be said that the present theory of statistical interpolation rests on an altogether satisfactory basis [except in evaluating mathematical functions]. The principles which govern it are not well defined, and the mathematical analysis of the methods, by which the principles should be brought into relation with the facts, is incomplete." In dealing with "accuracy," and the standard example of the index-number, much attention is given to "error" (not to be confused with the "law of

error" and the due importance of biased and unbiased error. It would not be just to Mr. Bowley to say that his conclusions are negative, but the impression of negative results is present.

Is it safe for one using statistics to assume that the figures before him have resulted from an unprejudiced or impartial investigation? By no means, even when they have been collected under the best apparent conditions tending to accuracy. In measuring *things* the standard taken as a measure may be wrong, and no number of separate measurements will eliminate the error. But knowing the degree of error an allowance may be made, just as one is made for the "personal equation" in scientific work. If then this liability to error exists in the case of things capable of definite measurement, how much greater must it be in the measuring, recording, or comparison of things offering no present standard for use, and for observing which only general directions can be given. The height of recruits, the chest measurements, the weight of an export or import, are comparatively definite qualities, of which some record approaching to accuracy may be made; but the social condition of a village, the wages received and retail prices, the value of imports, and changes in the value of the precious metals are difficult to measure in such a way as to leave no doubt of the correctness of the result. The methods to be applied in either case are the same, observation and tabulating; but in the one case a standard is at hand, in the other a standard must be found, and in groping for that standard the bias of the observers or their leaning to exaggerate or depreciate the facts before them must be discovered and due allowance made for the "error." In the end it is the "mean" or "average" which is the standard sought, but, unless the examination is restricted to a single quality or characteristic of the problem, confined to a locality of homogeneous conditions, or made under conditions of true and legitimate comparability, the error infects the conclusion.

Mr. Bowley lays down the important principle that "unbiased errors are of little importance compared with biased errors in a simple estimate; but biased errors diminish when the ratio of two similar estimates are taken." It is only a step from this to assert the rule that error may be lessened, even eliminated, by two processes: by *averaging*, which diminishes unbiased errors; and *comparison*, which diminishes biased errors. Thus far we seem to be resting upon firm ground. Our author adds:

When aiming at accuracy our principle always is to take care of the

pounds, and let the pence take care of themselves; and it is quite futile to diminish the unbiased errors, that is, to increase the precision of our measurements, while a large biased error runs through them all, If we do not know of the existence of biased errors, which in reality pervade our estimates, there is no remedy; if we do know of them, we are likely to obtain more accuracy by the most erroneous corrections for them than by neglecting them; for when we make unbiased corrections for our biased errors, we reduce them to unbiased errors, and then the more terms we include in our average the smaller is our resulting error.

This at first sight is more like an excuse for a free handling of statistics than a rule for attaining accuracy, for it would seem to justify the imposing a new bias upon an old one, thus increasing the error. In financial discussion in the United States the bias has been very marked, and no averages or comparisons, or efforts to correct the bias have been successful in ending doubt.

These examples, and the number could be extended, show that however useful in many directions, and however suggestive in others, the application of mathematics to statistics has proved, we are far from a "science" in the sense of absolute and universal rules. This is not to depreciate Mr. Bowley's volume, for he has performed a difficult task in a manner worthy of all praise. He has gathered the best results of those who have preceded him, and has given us a manual of method which marks the progress of this phase of statistical investigation. From this point of view no one interested in the subject can pass over the work, and its stimulating quality in obliging us to examine anew our principles and processes will advance the position of statistics many stages nearer that of a true science. The many diagrams and tables are interesting in themselves as well as illustrative of text and graphic methods. As a manual of method for testing statistics, and for introducing better arrangement and interpretation, it will be of high value, and may be taken in connection with such works as *Statistics and Sociology*, by Professor Mayo Smith, using the many examples in the latter as exercises in statistical method.

WORTHINGTON C. FORD.

BOSTON.

Seventh Annual Abstract of Labour Statistics of the United Kingdom. 1899–1900. London: Printed by Darling & Son, 1901. 8vo, pp. xv + 233.

ONE fact strikingly exhibited by this valuable compilation is the continued and increasing success of co-operative enterprises in Great